

STRUCTURE-PRESERVING MODEL REDUCTION OF FORCED HAMILTONIAN SYSTEMS*

LIQIAN PENG[†] AND KAMRAN MOHSENI[‡]

Abstract. This paper reports a development in the proper symplectic decomposition (PSD) for model reduction of forced Hamiltonian systems. As an analogy to the proper orthogonal decomposition (POD), PSD is designed to build a symplectic subspace to fit empirical data. Our aim is two-fold. First, to achieve computational savings for large-scale Hamiltonian systems with external forces. Second, to simultaneously preserve the symplectic structure and the forced structure of the original system. We first reformulate d'Alembert's principle in the Hamiltonian form. Corresponding to the integral and local forms of d'Alembert's principle, we propose two different structure-preserving model reduction approaches to reconstruct low-dimensional systems, based on the variational principle and on the structure-preserving projection, respectively. These two approaches are proven to yield the same reduced system. Moreover, by incorporating the vector field into the data ensemble, we provided several algorithms for energy preservation. In a special case when the external force is described by the Rayleigh dissipative function, the proposed method automatically preserves the dissipativity, boundedness, and stability of the original system. The stability, accuracy, and efficiency of the proposed method are illustrated through numerical simulations of a dissipative wave equation.

Key words. Structure-preserving, forced Hamiltonian systems, proper symplectic decomposition, d'Alembert's principle, variational principle, structure-preserving projection, dissipativity preservation

AMS subject classifications. 65P10, 37M15, 34C20, 93A15, 37J25

1. Introduction. For several centuries, physical models have been used in the natural sciences to describe and predict the world we live in. Physical models are anchored in venerated physical laws, such as Newton's laws of motion, Hamilton's principle, and conservation laws, to name but a few. Compared with data-based empirical models, physical models are more comprehensive and interpretable. In most cases, physical models derived from these laws are simple in the sense that they are typically expressed in terms of a few elegant equations. Nevertheless, for many practical problems, physical models become computational expensive, and even intractable, when they have high dimensions.

In recent years, the revolution in data sciences has opened a new window for understanding our world. Rather than discovering new physical laws, empirical models have led to powerful tools for extracting patterns and trends from the data directly. Although empirical models are phenomenological, they are predictive as well. When facing the curse of dimensionality in empirical models, dimensionality reduction techniques in data sciences provide low-cost solutions by pre-processing the data into a lower-dimensional form.

Can we couple tools from the data sciences—tools that are capable of dealing with high dimensions—with physical models? How can we combine the advantages of physical models with the information contained in the data? *Model reduction* is a technique for reducing the computational complexity of physical models in numerical simulations. Using empirical data, model reduction can provide low-dimensional

* The authors acknowledge partial support from the Air Force Office of Scientific Research (AFOSR).

[†]Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611-6250 (liqianpeng@ufl.edu).

[‡]Department of Mechanical and Aerospace Engineering, and Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6250 (mohseni@ufl.edu).

models that adequately predict the dynamics and allow for real-time analysis and control. The need for model reduction arises because, in many cases, direct numerical simulations are so computationally intensive that they either cannot be performed as often as needed or are only performed in special circumstances. See [1] for a survey on the classical model reduction methods.

Among these methods, the proper orthogonal decomposition (POD, also known as Karhunen-Loève decomposition or principle component analysis) with Galerkin projection, which was first introduced by Moore [16], has wide applications in many fields of science and engineering. As an empirical model reduction technique, the POD-Galerkin method (or POD for short) involves an offline-online splitting methodology. In the offline stage, empirical data is generated by experiments or direct numerical simulations. A reduced model (or reduced system) is then constructed by projecting the full model to a subspace where empirical data approximately resides. In the online stage, the reduced model is solved in the low-dimensional subspace.

However, the classical POD method is not guaranteed to yield a stable reduced model in general, even if the full model is stable [21, 20, 18]. The instability of a reduced model is often accompanied by blowup of system energy. Thus, a POD reduced model often fails to represent a physical system even if it is conservative or dissipative. More generally, POD can always yield a reduced model with a significantly lower dimension, but the reduced model might be merely a *numerical model*, rather than a *physical model*, since a POD reduced model may not yield to the underlying physical law that exists in the full model. Our primary motivation in this paper is to develop a model reduction technique such that the reduced model is guaranteed to be physical and as stable as the full model.

In the context of classical mechanics, d'Alembert's principle is the fundamental law of motion. Thus, if a reduced model remains physical, it should respect d'Alembert's principle. In section 2, we shall see that forced Hamiltonian equations satisfy d'Alembert's principle; conversely, if a system satisfies d'Alembert's principle, it can be represented by a forced Hamiltonian equation when choosing canonical coordinates. Thus, this paper focuses on developing a structure-preserving model reduction method for forced Hamiltonian systems, where the structure refers to the forced Hamiltonian structure and the systems are represented by ordinary differential equations (ODEs).

The proposed method in this paper extends our previous work on the symplectic model reduction of Hamiltonian systems to the structure-preserving model reduction of forced and dissipative systems. The symplectic model reduction is based on proper symplectic decomposition (PSD)-symplectic projection method [18]. Analogous to the POD-Galerkin method, PSD builds a symplectic subspace to fit empirical data, while the symplectic projection constructs a reduced Hamiltonian system on the symplectic subspace. Because the PSD reduced system preserves the symplectic structure, it automatically preserves the system energy and stability. Owing to these properties, the symplectic method outperforms the POD for model reduction of Hamiltonian systems, especially when stability is taken into consideration for long-time integration. Since many physical and engineering systems have external forces, this paper applies symplectic algorithms for model reduction of more general dynamical systems with external forces. Besides PSD, there are also other structure-preserving model reduction methods in the context of classical mechanics, including the Lagrangian approach [11, 4] and the port-Hamiltonian approach [7, 19, 5]. Compared with these methods, PSD is directly related to symplectic geometry, and provides more flexibility

to construct an optimal subspace to fit empirical data.

The remainder of this paper is organized as follows. Since forced Hamiltonian equations are anchored in d'Alembert's principle, we reformulate d'Alembert's principle in the Hamiltonian form in section 2. Corresponding to the integral and local forms of d'Alembert's principle, section 3 presents two different structure-preserving approaches for model reduction of forced Hamiltonian equations based on the variational principle and on the structure-preserving projection, respectively. We also prove that the two approaches are equal in the sense that they yield the same reduced equation and provide a PSD algorithm to construct the reduced basis function. In section 4, we discuss dissipative Hamiltonian systems, and prove that the proposed method automatically preserves the dissipativity. In section 5, the stability, accuracy, and efficiency of the proposed technique are illustrated through numerical simulations of a dissipative wave equation. Finally, conclusions are offered in section 6.

2. Forced Hamiltonian equations. In Hamiltonian mechanics, a mechanical system with external forces can be represented by a forced Hamiltonian equation. In this section, we first represent the forced Hamiltonian equation, and then derive it by d'Alembert's principle.

Let Q be an n -dimensional vector space over \mathbb{R} , Q^* be its dual space, and $\langle \cdot, \cdot \rangle : Q^* \times Q \rightarrow \mathbb{R}$ be a nondegenerate duality paring. With $\mathbb{V} = Q \oplus Q^*$, the pair (\mathbb{V}, Ω) is a symplectic vector space, where \mathbb{V} is the phase space and Ω is a closed non-degenerate two-form on \mathbb{V} . Assigning a symplectic form Ω to \mathbb{V} is referred to as giving \mathbb{V} a symplectic structure. With $u_q, v_q \in Q$ and $u_p, v_p \in Q^*$, we have

$$(2.1) \quad \Omega((u_q, u_p), (v_q, v_p)) = \langle v_p, u_q \rangle - \langle u_p, v_p \rangle.$$

Using canonical coordinates, Ω is represented by the Poisson matrix

$$J_{2n} = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix}.$$

Let $H : \mathbb{V} \rightarrow \mathbb{R}$ denote a smooth Hamiltonian function. The time evolution of forced Hamiltonian systems are defined by

$$(2.2) \quad \dot{q} = \nabla_p H(q, p), \quad \dot{p} = -\nabla_q H(q, p) + f_H(q, p),$$

where $q \in Q$ denotes the generalized coordinate, $p \in Q^*$ denotes the generalized momentum, and $f_H(q, p) \in Q^*$ is a force field. We abstract this formulation by introducing a variable $x = (q, p)$ in the phase space \mathbb{V} . Then, (2.2) becomes

$$(2.3) \quad \dot{x} = X_H(x) + X_F(x),$$

where $X_H(x) = J_{2n} \nabla_x H(x)$ denotes a Hamiltonian vector field, and $X_F(x) = (0, f_H(x))$ denotes a vertical vector field with zero in its first component. The state variable $x(t)$ can also be considered as a function of t , which gives a trajectory as t varies over \mathbb{R}_+ with a fixed initial condition x_0 . The trajectory $x(\mathbb{R}_+)$ contains a sequence of states that follow from x_0 .

Dissipative Hamiltonian systems are special forced Hamiltonian systems, where the system energy is decreasing with time. As an example, consider a one-dimensional harmonic oscillator with undamped angular frequency ω_0 and damping ratio ζ . Newton's second law takes the form

$$(2.4) \quad \ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2x = 0.$$

With $q = x$ and $p = \dot{x}$, the Hamiltonian function is given by $H(q, p) = \frac{1}{2}p^2 + \frac{1}{2}\omega_0^2 q^2$, and the force field is given by $f_H(q, p) = -2\zeta\omega_0 p$. Plugging $H(q, p)$ and $f_H(q, p)$ into (2.2), we can get the Hamiltonian representation of the harmonic oscillator, which is exactly the same as (2.4). The system energy is given by $E(t) := H(q(t), p(t))$. The time derivative of $E(t)$ is given by $\dot{E}(t) = -2\zeta\omega_0 p(t)^2$, which is negative for every t .

Forced Hamiltonian equations can be derived from the Legendre transformation of Euler–Lagrange equations with generalized forces. Alternatively, they can be directly obtained from the reformulation of d’Alembert’s principle in the Hamiltonian coordinates. In this paper, we take the second approach, since this approach also gives us insight on reconstructing structure-preserving reduced models. Our derivation closely follows reference [14] (pp. 205–210), where generalized Euler–Lagrangian equations are obtained from d’Alembert’s principle in the Lagrangian coordinates. We shall begin with the structure of the vertical vector field.

2.1. Force fields. Let $H : \mathbb{V} \rightarrow \mathbb{R}$ be a Hamiltonian function, $X_H : \mathbb{V} \rightarrow \mathbb{V}$ be the Hamiltonian vector field associated to H , and $\pi_Q : \mathbb{V} \rightarrow Q$, $(q, p) \mapsto q$ be the canonical projection. A vector field $X_F : \mathbb{V} \rightarrow \mathbb{V}$ is called *vertical* if the projection of X_F is zero, i.e., $\pi_Q(X_F) = 0$. Such a vector field X_F defines a one-form $\Delta^F : \mathbb{V} \rightarrow \mathbb{V}^*$ by contraction with Ω : $\Delta^F = -i_{X_F}\Omega$. For any vertical vector field u on \mathbb{V} , if the duality pairing $\Delta^F \cdot u = 0$, then we say Δ^F is *horizontal*. Here, we use the dot product to represent the duality pairing of \mathbb{V}^* and \mathbb{V} .

PROPOSITION 2.1. *If X_F is vertical, then the corresponding one-form Δ^F is horizontal. Conversely, given a horizontal one-form Δ^F on \mathbb{V} , the vector field X_F on \mathbb{V} , given by $\Delta^F = -i_{X_F}\Omega$, is vertical.*

Proof. Let $X_F = (f_q, f_p)$ and $u = (u_q, u_p)$, where $f_q, u_q \in Q$ and $f_p, u_p \in Q^*$. Using the definition of Ω , we have

$$(2.5) \quad \Delta^F \cdot u = -i_{X_F}\Omega(u) = -\Omega((f_q, f_p), (u_q, u_p)) = -\langle u_p, f_q \rangle + \langle f_p, u_q \rangle.$$

If u is vertical, $u_q = 0$. Thus, $\Delta^F \cdot u = 0$ for every vertical u is equivalent to $\langle u_p, f_q \rangle = 0$ for every u_p , which holds if and only if $f_q = 0$, i.e., the vector field X_F is vertical. \square

PROPOSITION 2.2. *A force field $f_H : \mathbb{V} \rightarrow Q^*$ induces a horizontal one-form Δ^F on \mathbb{V} by*

$$(2.6) \quad \Delta^F \cdot u = \langle f_H, \pi_Q(u) \rangle,$$

where u is a vector field on \mathbb{V} . Conversely, formula (2.6) defines a map f_H for any horizontal one-form Δ^F .

Proof. Given f_H , (2.6) defines a smooth one-form Δ^F on \mathbb{V} . If u is vertical, then the right-hand side of (2.6) vanishes, and so Δ^F is horizontal. Conversely, if Δ^F is a horizontal one-form on \mathbb{V} , $\Delta^F \cdot (u_q, u_p) = \Delta^F \cdot (u_q, 0) + \Delta^F \cdot (0, u_p) = \Delta^F \cdot (u_q, 0)$ for any vector field $u = (u_q, u_p)$. Thus, (2.6) is equivalent to $\langle f_H, u_q \rangle = \Delta^F \cdot (u_q, 0)$. Since $\langle \cdot, \cdot \rangle$ is nondegenerate, f_H is well-defined and smooth. \square

Propositions 2.1 and 2.2 imply that a force field f_H introduces a horizontal one-form Δ^F , which in turn determines a vertical vector field X_F . Using canonical coordinates, if $f_H(q, p) \in \mathbb{R}^n$ denotes a force field, then a horizontal one-form is given by $\Delta^F(q, p) = (f_H(q, p), 0)$. By contraction with Ω , the corresponding vertical field is given by $X_F(q, p) = -J_{2n}\Delta^F(q, p) = (0, f_H(q, p))$. Treating $\Delta^F(q, p)$ as the external force term on a mechanical system with a Hamiltonian H , we will next derive the equation of motion by d’Alembert’s principle.

2.2. D'Alembert's principle. D'Alembert's principle is a statement of the fundamental law of motion in classical mechanics. It is more general than Hamilton's principle since it considers both internal and external forces. In Newton's coordinates, the principle can be written as

$$(2.7) \quad \sum_i (F_i - \dot{p}_i) \cdot r_i = 0,$$

where F_i is the total applied force (excluding constraint forces) on the i -th particle, p_i is the momentum of the i -th particle, and δr_i is the virtual displacement of the i -th particle which is consistent with the constraints. We shall reformulate d'Alembert's principle by Hamiltonian coordinates.

DEFINITION 2.3. *Given a Hamiltonian function $H(q, p)$ and a horizontal one-form $\Delta^F(q, p)$, the integral d'Alembert's principle for a trajectory $(q(t), p(t))$ in \mathbb{V} is*

$$(2.8) \quad \delta \int_a^b L(q(t), p(t)) dt + \int_a^b \Delta^F(q(t), p(t)) \cdot (\delta q, \delta p) dt = 0,$$

where $(\delta q, \delta p)$ is a variation on \mathbb{V} , and $L : \mathbb{V} \rightarrow \mathbb{R}$ is the Lagrangian function defined by $L(q, p) = \langle p, \dot{q} \rangle - H(q, p)$.

The variation of the first term is given by the usual expression

$$\begin{aligned} \delta \int_a^b L(q, p) dt &= \int_a^b \langle p, \delta \dot{q} \rangle + \langle \delta p, \dot{q} \rangle - \langle \nabla_q H(q, p), \nabla_p H(q, p) \rangle \cdot (\delta q, \delta p) dt \\ &= \int_a^b \langle p, \delta \dot{q} \rangle + \langle \delta p, \dot{q} \rangle - \langle \nabla_q H(q, p), \delta q \rangle - \langle \delta p, \nabla_p H(q, p) \rangle dt \\ &= \int_a^b \langle -\dot{p} - \nabla_q H(q, p), \delta q \rangle + \langle \delta p, \dot{q} - \nabla_p H(q, p) \rangle dt \end{aligned}$$

for a given variation $(\delta q, \delta p)$, which vanishes at the endpoints. Since the external force $\Delta^F(q, p)$ is horizontal, Proposition (2.6), implies that

$$\Delta^F(q, p) \cdot (\delta q, \delta p) = \langle f_H(q, p), \delta q \rangle,$$

where f_H is the force field corresponding to Δ^F . Thus, (2.8) gives

$$(2.9) \quad \int_a^b \langle -\dot{p} - \nabla_q H(q, p) + f_H(q, p), \delta q \rangle + \langle \dot{q} - \nabla_p H(q, p), \delta p \rangle dt = 0.$$

Therefore, the trajectory of the integral d'Alembert's principle is given by the forced Hamiltonian equation (2.2).

We can also formulate an equivalent principle in terms of one-forms.

DEFINITION 2.4. *Given a Hamiltonian function $H(x)$ and a horizontal one-form $\Delta^F(x)$, the local d'Alembert's principle for the ultimate equation of motion, $\dot{x}(t) = X(x(t))$, is determined by*

$$(2.10) \quad i_X \Omega(x) = dH(x) - \Delta^F(x),$$

where X denotes the forced Hamiltonian vector field on \mathbb{V} .

PROPOSITION 2.5. *The two forms of d'Alembert's principle are equivalent, i.e., they give the same equation of motion.*

Proof. Let X_F denote the vector field associated to Δ^F , i.e., $\Delta^F = -i_{X_F}\Omega$. Since Δ^F is horizontal, X_F is vertical. Since $dH = i_{X_H}\Omega$,

$$(2.11) \quad X(x) = X_H(x) + X_F(x)$$

satisfies the local d'Alembert's principle. Conversely, the only vector field X satisfying the local d'Alembert's principle is given by (2.11), and uniqueness is guaranteed by nondegeneracy of Ω . Therefore, both the integral and local forms of d'Alembert's principle give the same vector field for the equation of motion $\dot{x}(t) = X(x(t))$. \square

From now on, we will refer to both (2.8) and (2.10) as simply d'Alembert's principle. By the above analysis, if a system satisfies d'Alembert's principle, the equation of motion is given by the forced Hamiltonian equation. Conversely, if a system is represented by a forced Hamiltonian equation, it automatically satisfies d'Alembert's principle. Since d'Alembert's principle is the first principle in classical mechanics, any mechanical system can be represented by a forced Hamiltonian equation. If a vector field X can be represented by the sum of a Hamiltonian vector field X_H and a vertical vector field X_F , we say X has forced Hamiltonian structure. In the next section, we develop a new model reduction method which preserves the forced Hamiltonian structure.

3. Reduction and reconstruction of dynamics. In this section, we propose two methods to construct reduced dynamics in a low-dimensional subspace. The first approach is based on the variational principle, which is closely related to the integral d'Alembert's principle; the second approach is based on the structure-preserving projection, which is closely related to the local d'Alembert's principle. Both methods take advantage of empirical data to construct a reduced system, while simultaneously preserving the underlying forced Hamiltonian structure. In other words, if the original system is a forced Hamiltonian equation, the reduced system remains a forced Hamiltonian equation, but with significantly fewer dimensions.

3.1. Variational principle. In the context of Lagrangian mechanics, the variational principle was used to yield reduced systems while preserving the Lagrangian structure [11, 4]. The idea is to insert $q = \Phi r$ into the Lagrangian to obtain a reduced system in terms of r . Here, $\Phi \in \mathbb{R}^{n \times k}$ denotes a POD basis matrix and $r \in \mathbb{R}^k$ denotes the reduced coordinates. Since \dot{q} is the time derivative of q , it is fixed by $\dot{q} = \Phi \dot{r}$. The Hamiltonian approach provides more flexibility, since q and p have the same status in the phase space.

Let (\mathbb{V}, Ω) and (\mathbb{W}, ω) be two symplectic vector spaces; $\dim(\mathbb{V}) = 2n$, $\dim(\mathbb{W}) = 2k$, and $k \leq n$. Using canonical coordinates, a lift $\sigma : \mathbb{V} \rightarrow \mathbb{W}, z \mapsto x$ can be written as

$$(3.1) \quad x = Az,$$

where $A \in \mathbb{R}^{2n \times 2k}$. Using the block form, $z = (r, s)$, $x = (q, p)$, and

$$(3.2) \quad A = \begin{bmatrix} A_{qq} & A_{qp} \\ A_{pq} & A_{pp} \end{bmatrix}.$$

Then, the map $x = \sigma(z)$ is represented by

$$\begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} A_{qq} & A_{qp} \\ A_{pq} & A_{pp} \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} A_{qq}r + A_{qp}s \\ A_{pq}r + A_{pp}s \end{bmatrix}.$$

In order to construct a reduced equation, we can plug $q = A_{qq}r + A_{qp}s$ and $p = A_{pq}r + A_{pp}s$ into (2.8) and take the variation on $(\delta r, \delta s)$. This yields

$$M \begin{bmatrix} \dot{r} \\ \dot{s} \end{bmatrix} = \begin{bmatrix} \nabla_r \tilde{H}(r, s) \\ \nabla_s \tilde{H}(r, s) \end{bmatrix} - \begin{bmatrix} A_{qq}^T \tilde{f}_H(r, s) \\ A_{qp}^T \tilde{f}_H(r, s) \end{bmatrix},$$

where $\tilde{H}(r, s) = H(A_{qq}r + A_{qp}s, A_{pq}r + A_{pp}s)$, $\tilde{f}_H(r, s) = f_H(A_{qq}r + A_{qp}s, A_{pq}r + A_{pp}s)$, and

$$M = \begin{bmatrix} A_{pq}^T A_{qq} - A_{qq}^T A_{pq} & A_{pq}^T A_{qp} - A_{qq}^T A_{pp} \\ A_{pp}^T A_{qq} - A_{qp}^T A_{pq} & A_{pp}^T A_{qp} - A_{qp}^T A_{pp} \end{bmatrix}.$$

Suppose M is invertible, we obtain

$$(3.3) \quad \begin{bmatrix} \dot{r} \\ \dot{s} \end{bmatrix} = M^{-1} \begin{bmatrix} \nabla_r \tilde{H}(r, s) \\ \nabla_s \tilde{H}(r, s) \end{bmatrix} - M^{-1} \begin{bmatrix} A_{qq}^T \tilde{f}_H(r, s) \\ A_{qp}^T \tilde{f}_H(r, s) \end{bmatrix}.$$

Equation (3.3) is the reduced system constructed by the variational principle. Next, we add some constraints to A such that (3.3) preserves the forced Hamiltonian structure.

DEFINITION 3.1. *Let (\mathbb{V}, Ω) and (\mathbb{W}, ω) be two symplectic vector spaces; $\dim(\mathbb{V}) = 2n$, $\dim(\mathbb{W}) = 2k$, and $k \leq n$. A lift $\sigma : \mathbb{W} \rightarrow \mathbb{V}$ is called symplectic if it preserves the symplectic structure:*

$$(3.4) \quad \omega(z, w) = \Omega(\sigma(z), \sigma(w)),$$

for every $z, w \in \mathbb{W}$.

Let A denote the matrix form of σ in canonical coordinates, then (3.4) implies

$$(3.5) \quad A^T J_{2n} A = J_{2k}.$$

In this case, we say the matrix A is symplectic, written as $A \in Sp(2k, \mathbb{R}^{2n})$, where

$$(3.6) \quad Sp(2k, \mathbb{R}^{2n}) := \{A \in \mathbb{R}^{2n \times 2k} \mid A^T J_{2n} A = J_{2k}\}$$

denotes the *symplectic Stiefel manifold*.

The symplectic condition can also be represented in block form by plugging (3.2) into (3.5).

PROPOSITION 3.2. *The matrix $A = [A_{qq}, A_{qp}; A_{pq}, A_{pp}]$ is symplectic if and only if $A_{qq}^T A_{pq}$ and $A_{qp}^T A_{pp}$ are symmetric and $A_{qq}^T A_{pp} - A_{pq}^T A_{qp} = I_k$.*

The next lemma gives a sufficient and necessary condition such that the variational approach is structure-preserving for any forced Hamiltonian equations.

LEMMA 3.3. *The reduced equation (3.3) constructed by the variational principle preserves the forced Hamiltonian structure for any Hamiltonian functions $H(q, p)$ and force fields $f_H(q, p)$ if and only if $A \in Sp(2k, \mathbb{R}^{2n})$ and $A_{qp} = 0$.*

Proof. If $A \in Sp(2k, \mathbb{R}^{2n})$ and $A_{qp} = 0$, then (3.3) reduces to

$$(3.7) \quad \begin{bmatrix} \dot{r} \\ \dot{s} \end{bmatrix} = J_{2n} \begin{bmatrix} \nabla_r \tilde{H}(r, s) \\ \nabla_s \tilde{H}(r, s) \end{bmatrix} + \begin{bmatrix} 0 \\ A_{qq}^T \tilde{f}_H(r, s) \end{bmatrix}.$$

where $\tilde{H}(r, s)$ represents the reduced Hamiltonian function and $A_{qp}^T \tilde{f}_H(r, s)$ represents the reduced force field.

Conversely, suppose that the reduced equation (3.3) preserves the forced Hamiltonian structure for any high-dimensional systems of the form (2.2). Let $f_H(q, p) = 0$, then (3.3) reduces to

$$(3.8) \quad \begin{bmatrix} \dot{r} \\ \dot{s} \end{bmatrix} = M^{-1} \begin{bmatrix} \nabla_r \tilde{H}(r, s) \\ \nabla_s \tilde{H}(r, s) \end{bmatrix}.$$

If this equation is Hamiltonian for any $\tilde{H}(r, s)$, we must have $M^{-1} = J_{2k}$. By Proposition 3.2, $M^{-1} = J_{2k}$ is equivalent to $A \in Sp(2k, \mathbb{R}^{2n})$. Now we plug $M^{-1} = J_{2k}$ into (3.3). If the second term on the right-hand side of (3.3) is a vertical, then $A_{qp}^T \tilde{f}_H(q, p) = 0$. Since $\tilde{f}_H(q, p)$ can be arbitrary, this implies that $A_{qp} = 0$. \square

3.2. Structure-preserving projection. In [18], the symplectic projection was proposed to construct reduced models for Hamiltonian equations while preserving the symplectic structure. In this section, we extend the symplectic projection to structure-preserving projection of forced Hamiltonian systems. The idea is to add some constraints to the symplectic projection so that the new projection also preserves the structure of the vertical vector field.

3.2.1. Symplectic projection. We begin with the basic definition of the symplectic projection.

DEFINITION 3.4. *Suppose $\sigma : \mathbb{W} \rightarrow \mathbb{V}$ is a symplectic lift. Then the adjoint of σ is the linear mapping $\pi : \mathbb{V} \rightarrow \mathbb{W}$ satisfying*

$$(3.9) \quad \omega(w, \pi(x)) = \Omega(\sigma(w), x),$$

for every $w \in \mathbb{W}$ and $x \in \mathbb{V}$. We say π is the symplectic projection induced by σ .

Using canonical coordinates, σ can be represented by a symplectic matrix A . Then, the symplectic projection $\pi : x \mapsto z$ can be written as

$$(3.10) \quad z = A^+ x,$$

where $A^+ \in \mathbb{R}^{2k \times 2n}$. Equation (3.9) implies that $J_{2k} A^+ = A^T J_{2n}$. Since J_{2k} is invertible, it follows that

$$(3.11) \quad A^+ = J_{2k}^T A^T J_{2n}.$$

Since $A^T J_{2n} A = J_{2k}$, A^+ is a left inverse of A , i.e.,

$$(3.12) \quad A^+ A = I_{2k}.$$

In general, A^+ is not equal to the Moore–Penrose pseudoinverse $(A^T A)^{-1} A^T$, and the left inverse of A is not unique. However, since Ω and ω are nondegenerate, A^+ is the unique adjoint matrix of A with respect to the Poisson matrices J_{2n} and J_{2k} .

Equation (3.12) implies that $\pi \circ \sigma = id_{\mathbb{W}}$. Since $(\sigma \circ \pi) \circ (\sigma \circ \pi) = \sigma \circ (\pi \circ \sigma) \circ \pi = \sigma \circ \pi$, $\sigma \circ \pi$ defines a projection operator on \mathbb{V} .

PROPOSITION 3.5. *Suppose $\sigma : \mathbb{W} \rightarrow \mathbb{V}$ is a symplectic lift and $\pi : \mathbb{V} \rightarrow \mathbb{W}$ is a symplectic projection introduced by σ . Then*

$$(3.13) \quad \Omega(u, (\sigma \circ \pi)(v)) = \Omega((\sigma \circ \pi)(u), v) = \Omega((\sigma \circ \pi)(u), (\sigma \circ \pi)(v)) = \omega(\pi(u), \pi(v)),$$

for every $u, v \in \mathbb{V}$.

Proof. Using canonical coordinates, (3.13) can be rewritten as

$$(3.14) \quad u^T J_{2n} A A^+ v = (A A^+ u)^T J_{2n} v = (A A^+ u)^T J_{2n} (A A^+ v) = (A^+ u)^T J_{2k} (A^+ v),$$

which can be verified by replacing A^+ with (3.11). \square

Since Ω is skew-symmetric, Proposition 3.5 implies that

$$\Omega(u, (\sigma \circ \pi)(u)) = \Omega((\sigma \circ \pi)(u), u) = 0.$$

As a consequence, for every $u, v \in \mathbb{V}$,

$$(3.15) \quad \Omega(u, (\sigma \circ \pi)(u + v)) = \Omega(u, (\sigma \circ \pi)(v)).$$

The symplectic projection defines a mapping from a high-dimensional space to a low-dimensional space. The same projection can also be applied a high-dimensional Hamiltonian system to obtain a reduced system while preserving the symplectic structure.

To see this, suppose the original system is Hamiltonian, i.e., $\dot{x} = X_H(x)$. Suppose $x = Az$. Using the chain rule, we obtain $\nabla_z H(Az) = A^T \nabla_x H(x)$. Using $A^+ J_{2n} = J_{2k} A^T$, we obtain the symplectic projection of the tangent vector at x ,

$$(3.16) \quad \pi(X_H(x)) = A^+ J_{2n} \nabla_x H(x) = J_{2k} A^T \nabla_x H(x) = J_{2k} \nabla_z H(Az) = J_{2k} \nabla_z \tilde{H}(z),$$

where $\tilde{H}(z) := H(Az)$ defines a Hamiltonian function on \mathbb{W} . Since $X_{\tilde{H}}(z) := J_{2k} \nabla_z \tilde{H}(z)$ gives a Hamiltonian vector field on \mathbb{W} , the reduced system $\dot{z} = X_{\tilde{H}}(z)$ is a well-defined and preserves the symplectic structure.

With some extra constraints, the next section shows that the symplectic projection can also be applied to a forced Hamiltonian system to construct a reduced system while preserving the structure of the vertical vector field.

3.2.2. Structure-preserving projection. For a forced Hamiltonian system, the corresponding vector field is given by $X(x) = X_H(x) + X_F(x)$ at each $x \in \mathbb{V}$. Then, we can define a reduced vector field by

$$(3.17) \quad \pi(X(\sigma(z))) = \pi(X_H(\sigma(z))) + \pi(X_F(\sigma(z))),$$

at each $z \in \mathbb{W}$. If σ and π respectively represent the symplectic lift and symplectic projection, the last section shows that $\pi(X_H(\sigma(z)))$ gives a Hamiltonian vector field on \mathbb{W} . Thus, if the reduced system preserves the forced Hamiltonian structure, we only need $\pi(X_F(\sigma(z)))$ to be a vertical vector field on \mathbb{W} .

In block form, A^+ can be written as

$$(3.18) \quad A^+ = \begin{bmatrix} A_{pp}^T & -A_{qp}^T \\ -A_{pq}^T & A_{qq}^T \end{bmatrix}.$$

It follows that the projection $X_{\tilde{F}}(z)$ of the vertical vector field X_F at Az is given by

$$(3.19) \quad X_{\tilde{F}}(z) = A^+ X_F(Az) = \begin{bmatrix} A_{pp}^T & -A_{qp}^T \\ -A_{pq}^T & A_{qq}^T \end{bmatrix} \begin{bmatrix} 0 \\ f_H(Az) \end{bmatrix} = \begin{bmatrix} -A_{qp}^T f_H(Az) \\ A_{qq}^T f_H(Az) \end{bmatrix}.$$

Thus, the symplectic projection preserves the forced structure, i.e., $\pi X_F(Az)$ is vertical if and only if $A_{qp}^T f_H(Az) = 0$ for any $f_H(Az)$. This is equivalent to $A_{qp} = 0$.

DEFINITION 3.6. Let $z \in \mathbb{W}$ and $x \in \mathbb{V}$. Using the canonical coordinates, a linear mapping $\pi : x \mapsto z$ is a structure-preserving projection if there exists a symplectic matrix $A \in Sp(2k, \mathbb{R}^{2n})$ with $A_{qp} = 0$, such that

$$(3.20) \quad z = A^+x.$$

Now, suppose $A \in Sp(2k, \mathbb{R}^{2n})$, $A_{qp} = 0$, and $x(t) \in \text{Range}(A)$ for every t . Then, $x(t) = Az(t)$. Taking the time derivative of $z = A^+x$ and using (2.3), the time evolution of $z(t)$ is given by

$$(3.21) \quad \dot{z} = \pi \dot{x} = \pi X_H(x) + \pi X_F(x) = X_{\tilde{H}}(z) + X_{\tilde{F}}(z).$$

Even if $x(t) \notin \text{Range}(A)$ for some t , the last expression is still well-defined forced Hamiltonian vector field. Thus, the reduced system constructed by the structure-preserving projection preserves the forced Hamiltonian structure.

REMARK 3.7. Both the variational principle and the structure-preserving projection methods requires that $A \in Sp(2k, \mathbb{R}^{2n})$ and $A_{qp} = 0$, which means that the two methods can share the same basis matrix A . Let $z = (r, s)$, then (3.7) and (3.21) define the same system on the subspace spanned by the column vectors of A . Thus, two methods construct the same reduced system. From now on, we do not distinguish the variational principle and the structure-preserving projection when we mention a structure-preserving reduced model.

DEFINITION 3.8. Given a $2n$ -dimensional forced Hamiltonian system (2.3) with an initial condition $x_0 \in \mathbb{R}^{2n}$, the structure-preserving reduced model is a $2k$ -dimensional ($k \leq n$) system

$$(3.22) \quad \dot{z} = X_{\tilde{H}}(z) + X_{\tilde{F}}(z),$$

with the initial condition $z_0 = A^+x_0 \in \mathbb{R}^{2k}$, where $A \in Sp(2k, \mathbb{R}^{2n})$, $A_{qp} = 0$, $X_{\tilde{H}}(z)$ and $X_{\tilde{F}}(z)$ respectively represent the Hamiltonian vector field and the vertical vector on \mathbb{R}^{2k} .

REMARK 3.9. The vertical vector field $X_{\tilde{F}}$ defines a horizontal one-form $\Delta^{\tilde{F}}$ by contraction with ω , i.e., $\Delta^{\tilde{F}} = -i_{X_{\tilde{F}}}\omega$. It follows that

$$(3.23) \quad i_Z\omega = d\tilde{H} - \Delta^{\tilde{F}}.$$

This verifies that the reduced system constructed by the structure-preserving projection also satisfies the local d'Alembert's principle.

REMARK 3.10. Suppose π is a structure-preserving projection. Then, there exists a symplectic matrix $A \in Sp(2k, \mathbb{R}^{2n})$ with $A_{qp} = 0$. In the block form, this implies that

$$(3.24) \quad A = \begin{bmatrix} A_{qq} & 0 \\ A_{pq} & A_{pp} \end{bmatrix},$$

$A_{qq}^T A_{pq}$ is symmetric, and $A_{qq}^T A_{pp} = I_k$. Using (3.24), the projection operator $\sigma \circ \pi$ has the form

$$(3.25) \quad AA^+ = \begin{bmatrix} A_{qq}A_{pp}^T & 0 \\ A_{pq}A_{pp}^T - A_{pp}A_{pq}^T & A_{pp}A_{qq}^T \end{bmatrix},$$

which gives an invariant subspace, $0 \oplus Q^*$, of \mathbb{V} . Thus, all vectors $u \in 0 \oplus Q^*$ are transformed by $\sigma \circ \pi$ into vectors that are also contained in $0 \oplus Q^*$. This can be stated as

$$(3.26) \quad u \in 0 \oplus Q^* \Rightarrow (\sigma \circ \pi)(u) \in 0 \oplus Q^*.$$

3.3. Proper symplectic decomposition (PSD). PSD is an empirical model reduction technique, where the empirical data is used to construct a symplectic basis matrix A . Let $q(t_i), p(t_i) \in \mathbb{R}^n$ ($i = 1, \dots, N$) denote the empirical data. Assume $n \geq 2N$. Rewriting the state variable in the form $x(t_i) = [q(t_i); p(t_i)]$, we can define a snapshot matrix in $\mathbb{R}^{2n \times N}$,

$$(3.27) \quad M_x := [x(t_1), \dots, x(t_N)].$$

The structure-preserving projection of M_x onto a low dimensional subspace is given by $M_z = A^+ M_x$, where $A \in Sp(2k, \mathbb{R}^{2n})$, $M_z = [z(t_1), \dots, z(t_N)] \in \mathbb{R}^{2k \times N}$, and $z(t_i) = A^+ x(t_i)$. The same projection of M_x in the original coordinates is given by AM_z , or $AA^+ M_x$.

The Frobenius norm $\|\cdot\|_F$ can be used to measure the error between M_x and its projection AM_z . Suppose a symplectic matrix A minimizes the projection error in a least squares sense. Then, A is a solution of the following optimization problem:

$$(3.28) \quad \begin{aligned} & \text{minimize} \quad \|M_x - AA^+ M_x\|_F \\ & \text{subject to} \quad A^T J_{2n} A = J_{2k} \quad \text{and} \quad A_{qp} = 0. \end{aligned}$$

Let $M_q = [q(t_1), \dots, q(t_N)]$ and $M_p = [p(t_1), \dots, p(t_N)]$. Using (3.25), the cost function in (3.28) can be expanded as

$$\|M_q - A_{qq} A_{pp}^T M_q\|_F + \|M_p - (A_{pq} A_{pp}^T - A_{pp} A_{pq}^T) M_q - A_{pp} A_{qq}^T M_p\|_F.$$

By Remark 3.7, the constraint in (3.28) holds if and only if $A_{qq}^T A_{pq}$ is symmetric and $A_{qq}^T A_{pp} = I_k$. Thus, (3.28) is equivalent to

$$(3.29) \quad \begin{aligned} & \text{minimize} \quad \|M_q - A_{qq} A_{pp}^T M_q\|_F + \|M_p - (A_{pq} A_{pp}^T - A_{pp} A_{pq}^T) M_q - A_{pp} A_{qq}^T M_p\|_F \\ & \text{subject to} \quad A_{qq}^T A_{pq} = A_{pq}^T A_{qq} \quad \text{and} \quad A_{qq}^T A_{pp} = I_k. \end{aligned}$$

Since matrices A_{qq} , A_{pq} and A_{qp} all have $n \times k$ elements and the constraint region is nonconvex, it is expected to be quite expensive to solve (3.29) by nonconvex nonlinear programming. To this effect, we use a singular value decomposition (SVD)-based method, *cotangent lift*, to construct a near optimal symplectic matrix in a subset of $Sp(2k, \mathbb{R}^{2n})$ with $A_{qp} = 0$. The idea is to search for the optimal matrix, A_1 , in a subset of $Sp(2k, \mathbb{R}^{2n})$ with $A_{qp} = 0$, such that all the empirical data lies near $\text{Range}(A_1)$. In particular, we assume that

$$(3.30) \quad A_{qp} = A_{pq} = 0 \quad \text{and} \quad A_{qq} = A_{qq} = \Phi.$$

Then, $A_1 = \text{diag}(\Phi, \Phi)$ for some $\Phi \in \mathbb{R}^{n \times k}$. It is straightforward to verify that $A_1^T J_{2n} A_1 = J_{2k}$ if and only if $\Phi^T \Phi = I_k$. Under these assumptions, (3.29) reduces to

$$(3.31) \quad \begin{aligned} & \text{minimize} \quad \|M_q - \Phi \Phi^T M_q\|_F + \|P - \Phi \Phi^T P\|_F \\ & \text{subject to} \quad \Phi^T \Phi = I_k. \end{aligned}$$

The cost function in (3.31) equals $\|M_{q,p} - \Phi\Phi^T M_{q,p}\|_F$, where

$$(3.32) \quad M_{q,p} := [q(t_1), \dots, q(t_N), p(t_1), \dots, p(t_N)]$$

defines an extended snapshot matrix $M_{q,p} \in \mathbb{R}^{n \times 2N}$.

Algorithm 1 Cotangent Lift

Require: An empirical data ensemble $\{q(t_i), p(t_i)\}_{i=1}^N$.

Ensure: A symplectic matrix A_1 in block-diagonal form.

- 1: Construct an extended snapshot matrix $M_{q,p}$ as (3.32).
 - 2: Compute the SVD of $M_{q,p}$ to obtain a POD basis matrix Φ .
 - 3: Construct the symplectic matrix $A_1 = \text{diag}(\Phi, \Phi)$.
-

If Φ^* denotes the optimal value of Φ in (3.31), Φ^* can be directly solved by the SVD of $M_{q,p}$. Thus, the cotangent lift method simplifies the optimization problem (3.28) to an SVD problem. Algorithm 1 lists the detailed procedure of the cotangent lift method. Since the SVD of a $n \times m$ ($n \geq m$) matrix requires $2nm^2 + 2m^3$ operations [22], the computational cost of Algorithm (1) is $8nN^2 + 16N^3$, which is linearly dependent on n .

The cost function in (3.31) is also equal to the projection error of the empirical data in the Frobenius norm $\|M_x - AA^+ M_x\|_F$. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{2N} \geq 0$ denote the singular values of $M_{q,p}$ in decreasing order. Then, the projection error of the cotangent lift method is determined by the truncated singular values of $M_{q,p}$,

$$(3.33) \quad E_{\text{COT}}^{(2k)} = \left\| (I - \Phi^* \Phi^{*T}) M_{q,p} \right\|_F = \sqrt{\sum_{i=k+1}^{2N} \lambda_i^2},$$

where we use the superscript $2k$ to emphasize that the symplectic subspace spanned by the column vectors of $A_1 = \text{diag}(\Phi, \Phi)$ has dimension $2k$.

PROPOSITION 3.11. *Let $E_{\text{COT}}^{(2k)}$, $E_{\text{POD}}^{(2k)}$, and $E_{\text{OPT}}^{(2k)}$ denote the projection error by the cotangent lift method, POD, and the nonlinear programming method to solve (3.28), respectively. Then,*

$$(3.34) \quad \frac{1}{2} E_{\text{cot}}^{(4k)} \leq E_{\text{POD}}^{(2k)} \leq E_{\text{OPT}}^{(2k)} \leq E_{\text{COT}}^{(2k)}.$$

Proof. The cotangent lift method yields an optimal symplectic matrix in $\mathbb{M}_1 := \{A \in \mathbb{R}^{2n \times 2k} : A = \text{diag}(\Phi, \Phi), \Phi \in \mathbb{R}^{n \times k}, \text{ and } \Phi^T \Phi = I_k\}$. The feasible set of the optimization problem (3.28) is given by $\mathbb{M} := \{A \in \mathbb{R}^{2n \times 2k} : AJ_{2n}A = J_{2k} \text{ and } A_{qp} = 0\}$. POD can find the most optimal matrix in $\mathbb{R}^{2n \times 2k}$ to minimize the projection error. Thus, corresponding to $\mathbb{M}_1 \subset \mathbb{M} \subset \mathbb{R}^{2n \times 2k}$, we have $E_{\text{POD}}^{(2k)} \leq E_{\text{OPT}}^{(2k)} \leq E_{\text{COT}}^{(2k)}$.

Next we will prove $\frac{1}{2} E_{\text{cot}}^{(4k)} \leq E_{\text{POD}}^{(2k)}$. According to (3.33), $(E_{\text{cot}}^{(4k)})^2 = \sum_{i=2k+1}^{2N} \lambda_i^2$. Similar, if $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_N \geq 0$ denotes all the singular values of M_x in descending order, then $(E_{\text{POD}}^{(2k)})^2 = \sum_{i=k+1}^N \sigma_i^2$.

Since $M_{q,p} = \begin{bmatrix} Q \\ P \end{bmatrix}$, $M_{q,p}^T M_{q,p} = \begin{bmatrix} Q^T Q & Q^T P \\ P^T Q & P^T P \end{bmatrix} \in \mathbb{R}^{2N \times 2N}$, with eigenvalues $\{\lambda_i^2\}_{i=1}^{2N}$. Since $M_x = \begin{bmatrix} Q & P \end{bmatrix}$, $M_x^T M_x = Q^T Q + P^T P \in \mathbb{R}^{N \times N}$, with eigenvalues $\{\sigma_i^2\}_{i=1}^N$. By the construction, both $M_{q,p}^T M_{q,p}$ and $M_x^T M_x$ are positive-semidefinite.

Let $S = \text{diag}\{M_x^T M_x, M_x^T M_x\}$, then S is also positive-semidefinite, and the i th largest eigenvalue of S is $(\sigma_{\lceil \frac{i}{2} \rceil})^2$. Moreover,

$$\begin{aligned} 2S - M_{q,p}^T M_{q,p} &= \begin{bmatrix} 2M_x^T M_x & 0 \\ 0 & 2M_x^T M_x \end{bmatrix} - \begin{bmatrix} Q^T Q & Q^T P \\ P^T Q & P^T P \end{bmatrix} \\ &= \begin{bmatrix} Q^T Q & -Q^T P \\ -P^T Q & P^T P \end{bmatrix} + 2 \begin{bmatrix} P^T P & 0 \\ 0 & P^T P \end{bmatrix} \\ &= \begin{bmatrix} Q^T \\ -P^T \end{bmatrix} [Q \quad -P] + 2 \begin{bmatrix} P^T P & 0 \\ 0 & P^T P \end{bmatrix}. \end{aligned}$$

The last equation implies that $2S \geq M_{q,p}^T M_{q,p} \geq 0$. By the min-max theorem, the i th largest eigenvalue of $2S$ is greater than the i th largest eigenvalue of $M_{q,p}^T M_{q,p}$. This implies that $2(\sigma_{\lceil \frac{i}{2} \rceil})^2 \geq \lambda_i^2$. It follows that

$$(E_{\text{cot}}^{(4k)})^2 = \sum_{i=2k+1}^{2N} \lambda_i^2 \leq 2 \sum_{i=2k+1}^{2N} (\sigma_{\lceil \frac{i}{2} \rceil})^2 = 4 \sum_{i=k+1}^N \sigma_i^2 = 4(E_{\text{POD}}^{(2k)})^2.$$

This completes the proof. \square

3.4. Energy preservation. Let $x(t)$ be the solution of (2.3) with $x(0) = x_0$, and $E(t) = H(x(t))$ be the corresponding system energy at time t . Since $dH(x) \cdot X_H(x) = \Omega(X_H(x), X_H(x)) = 0$, the time derivative of $E(t)$ equals

$$(3.35) \quad \dot{E}(t) = dH \cdot X|_{x(t)} = dH \cdot X_F|_{x(t)}.$$

Thus, the time derivative of system energy is completely determined by the Hamiltonian function $H(x)$ and the vertical vector field $X_F(x)$.

PROPOSITION 3.12. *The time derivative of $E(t)$ can also be represented in terms of the force field, i.e.*

$$(3.36) \quad dH \cdot X_F|_x = \langle f_H, \dot{q} \rangle|_x,$$

where $\dot{q}(x) = \pi_Q(X_H(x)) = \nabla_p H(q, p)$.

Proof. Let X_F be a vertical vector field. By Proposition 2.1, X_F induces a horizontal one-form $\Delta^F = -i_{X_F} \Omega$ on \mathbb{V} ; by Proposition 2.2, Δ^F in turn induces a force field f_H , which is given by

$$(3.37) \quad \langle f_H, \pi_Q(u) \rangle = \Delta^F \cdot u = -\Omega(X_F, u),$$

where u is a vector field on \mathbb{V} . Let X_H denote the Hamiltonian vector field. Then,

$$dH \cdot X_F = (i_{X_H} \Omega) \cdot X_F = \Omega(X_H, X_F) = -\Omega(X_F, X_H) = \langle f_H, \pi_Q(X_H) \rangle,$$

which gives (3.36). \square

Let $\tilde{E}(t) = \tilde{H}(z(t))$ denote the system energy of the forced Hamiltonian system (3.22) in reduced coordinates. Similar to (3.35), the time derivative of $\tilde{E}(t)$ is given by

$$(3.38) \quad \dot{\tilde{E}}(t) = d\tilde{H} \cdot X_{\tilde{F}}|_{z(t)}.$$

THEOREM 3.13. *The reduced forced Hamiltonian system exactly preserves the time derivative of system energy at $z \in \mathbb{W}$, i.e. $d\tilde{H} \cdot X_{\tilde{F}}|_z = dH \cdot X_F|_{\sigma(z)}$, if any one of the following conditions is satisfied at $\sigma(z)$:*

- (a) $(\sigma \circ \pi)(X_F) = X_F$.
- (b) $(\sigma \circ \pi)(X_H) = X_H$.
- (c) $(\sigma \circ \pi)(X_H) = X$.
- (d) $(\pi_P \circ \sigma \circ \pi)(X_F) = \pi_P(X_F)$.
- (e) $(\pi_Q \circ \sigma \circ \pi)(X_H) = \pi_Q(X_H)$.

Proof. Using (3.13), we obtain

$$\begin{aligned}
 d\tilde{H} \cdot X_{\tilde{F}}|_z &= \omega(X_{\tilde{H}}, X_{\tilde{F}})|_z = \omega(\pi(X_H(\sigma(z))), \pi(X_F(\sigma(z)))) \\
 (3.39) \quad &= \Omega(X_H, (\sigma \circ \pi)X_F)|_{\sigma(z)} \\
 &= \Omega((\sigma \circ \pi)X_H, X_F)|_{\sigma(z)}.
 \end{aligned}$$

If (a) or (b) holds, the second or third line of (3.39) would imply that $d\tilde{H} \cdot X_{\tilde{F}}|_z = \Omega(X_H, X_F)|_{\sigma(z)}$, which equals $dH \cdot X_F|_{\sigma(z)}$.

Using (3.15), the time derivative of system energy for the reduced system can be represented by

$$(3.40) \quad d\tilde{H} \cdot X_{\tilde{F}}|_z = \Omega(X_H, (\sigma \circ \pi)(X_H + X_F))|_{\sigma(z)} = \Omega(X_H, (\sigma \circ \pi)(X))|_{\sigma(z)}.$$

This implies that (c) is a sufficient condition to preserves the time derivative of system energy.

Since both X_F and $(\sigma \circ \pi)X_F$ are vertical vector fields, we have $\pi_Q((\sigma \circ \pi)X_F) = \pi_Q(X_F) = 0$. Thus, (a) is equivalent to (d).

Finally, by (3.39) and (3.37), we obtain

$$\begin{aligned}
 d\tilde{H} \cdot X_{\tilde{F}}|_z &= \Omega((\sigma \circ \pi)(X_H), X_F)|_{\sigma(z)} = -\Omega(X_F, (\sigma \circ \pi)(X_H))|_{\sigma(z)} \\
 &= \langle f_H, \pi_Q((\sigma \circ \pi)(X_H)) \rangle|_{\sigma(z)}.
 \end{aligned}$$

Thus, if (e) is satisfied, the time derivative of system energy is preserved as well. \square

3.4.1. Optimization of the basis matrix. Theorem 3.13 implies that the time derivative of system energy is exactly preserved when the vector fields X_H , X_F , X , or their vertical/horizontal components are invariant under the projection operator $\sigma \circ \pi$. Motivated by this, we formulate five optimization problems corresponding to each individual condition in Theorem 3.13. All the optimization problems seek to construct a symplectic basis matrix A such that one of the aforementioned vector fields (or their vertical/horizontal components) can lie near the subspace spanned by the column vectors of A .

Condition (a). The condition (a) can be written as $AA^+X_F = X_F$ in canonical coordinates, which requires that $X_F(Az) \in \text{Range}(A)$ for each z . To satisfy this condition approximately, we can construct an extended data ensemble,

$$(3.41) \quad M_{x, X_F} := [x(t_1), \dots, x(t_N), X_F(x(t_1)), \dots, X_F(x(t_N))],$$

and then construct a symplectic matrix A to fit each column vector of M_{x, X_F} by solving the following optimization problem:

$$\begin{aligned}
 (3.42) \quad &\text{minimize} \quad \|M_{x, X_F} - AA^+M_{x, X_F}\|_F \\
 &\text{subject to} \quad A^T J_{2n} A = J_{2k} \quad \text{and} \quad A_{qp} = 0.
 \end{aligned}$$

Condition (b). Replacing X_F (resp., M_{x,X_F}) with X_H (resp., M_{x,X_H}) in (3.41) and solving (3.42) to minimize M_{x,X_H} will yield a symplectic matrix A to satisfy (b) approximately.

Condition (c). If we replace X_F with X in (3.41), then (c) will be approximately satisfied by the similar procedure. In this case, the symplectic matrix A is constructed to fit both the solution snapshots x and time derivative X of x simultaneously to preserve the time derivative of system energy. In previous literature, the analogous idea of incorporating time derivative snapshots [3, 17] or difference quotients [9, 10, 8], into the data ensemble has been widely used to enhance the performance (such as convergence and accuracy) of POD reduced models.

Condition (d). In order to approximately satisfy (d), we first construct a data ensemble in $\mathbb{R}^{n \times N}$ for the force field

$$(3.43) \quad M_{f_H} := [f_H(x(t_1)), \dots, f_H(x(t_N))].$$

Since the vertical vector field $X_F(x)$ can be represented by $X_F(x) = [0; f_H(x)]$, the corresponding data ensemble for the vertical vector field is given by

$$(3.44) \quad M_{X_F} := \begin{bmatrix} 0 \\ M_{f_H} \end{bmatrix},$$

where $M_{X_F} \in \mathbb{R}^{2n \times N}$. Thus, an optimal value of A can be obtained by minimizing a cost function that is related to the projection error of x and $\pi_P(X_F(x))$. In particular, the cost function can be formulated as

$$(3.45) \quad \|M_x - AA^+M_x\|_F + \gamma \|I_P M_{X_F} - I_P AA^+M_{X_F}\|_F,$$

where $I_P = [0, I_n] \in \mathbb{R}^{n \times 2n}$ is the matrix representation of π_P , and γ is a weighting coefficient to balance the truncation of M_x and $I_P M_{X_F}$. Replacing AA^+ by (3.25) simplifies the cost function (3.45) to

$$(3.46) \quad \|M_x - AA^+M_x\|_F + \gamma \|M_{f_H} - A_{pp}A_{qq}^T M_{f_H}\|_F.$$

When $\gamma = 1$, the cost functions in (3.42) and (3.46) are exactly the same.

Condition (e). In order to satisfy (e) approximately, one can construct a data ensemble that contains $\nabla_p H(x)$, and form an optimization function in terms of $I_Q = [I_n, 0] \in \mathbb{R}^{n \times 2n}$.

While nonconvex nonlinear programming can result in the most optimal subspace to fit an extended data ensemble, considering A is a matrix with $2n \times 2k$ elements, the programming problem can be very expensive and even intractable. Thus, we shall propose a cotangent lift method to obtain a near optimal value of A at a relatively lower cost while simultaneously preserving the time derivative of system energy.

3.4.2. Cotangent lift with energy preservation. The cotangent lift method can simplify all the optimization problems mentioned in the previous section. As an example, we shall give a cotangent lift algorithm to minimize the cost function (3.46) with $\gamma = 1$, such that (a) and (d) can be satisfied.

The cotangent lift methods requires that $A_{qp} = A_{pq} = \Phi$ and $A_{qq} = A_{pp} = \Phi$, where Φ is an orthonormal matrix. Then, q , p , and f_H have the same status in the cost function, and all the data of q , p , and f_H should lie near the Range of A . As a result, the k columns of Φ can be obtained from the left singular vectors of the following data ensemble

$$(3.47) \quad M_{q,p,f_H} := [q(t_1), \dots, q(t_N), p(t_1), \dots, p(t_N), f_H(x(t_1)), \dots, f_H(x(t_N))].$$

Algorithm 2 Cotangent lift with energy preservation**Require:** An empirical data ensemble $\{q(t_i), p(t_i), f_H(x(t_i))\}_{i=1}^N$.**Ensure:** A symplectic matrix A_1 in block-diagonal form.

- 1: Construct an extended snapshot matrix M_{q,p,f_H} as (3.47).
- 2: Compute the SVD of M_{q,p,f_H} to obtain a POD basis matrix Φ .
- 3: Construct the symplectic matrix $A_1 = \text{diag}(\Phi, \Phi)$.

Algorithm 2 lists the detailed procedure for the cotangent lift method for the preservation of the time derivative of system energy. Since this algorithm is based on SVD, the reconstruction error of the time derivative of system energy at $x = \sigma(z)$ can be estimated by the following:

$$\begin{aligned}
\|dH \cdot X_F|_x - d\tilde{H} \cdot X_{\tilde{F}}|_z\| &= \|\Omega(X_H, X_F)|_x - \Omega(X_H, (\sigma \circ \pi)(X_F))|_x\| \\
&\leq \|X_H(x)\| \cdot \|X_F(x) - AA^+X_F(x)\| \\
&= \|X_H(x)\| \cdot \|f_H(x) - A_{pp}A_{qq}^T f_H(x)\|.
\end{aligned}$$

In a compact subset M of \mathbb{V} , we can assume $\|X_H(x)\|$ to be uniformly bounded. If the data set of $f_H(x)$ is representative at the solution trajectory, $\|f_H(x) - A_{pp}A_{qq}^T f_H(x)\|$ is bounded by a constant multiplied by the truncated singular values of (3.47).

4. Reduction of dissipative Hamiltonian systems. In this section, we discuss a special form of forced Hamiltonian systems where the vertical vector field is dissipative. We also prove that the proposed model reduction method preserves the stability of the dissipative Hamiltonian system.

4.1. Dissipative Hamiltonian systems. We begin with the definition of a dissipative vector field.

DEFINITION 4.1. *A vertical vector field X_F on \mathbb{V} is called dissipative if $dH \cdot X_F|_x \leq 0$ for every $x \in \mathbb{V}$.*

Proposition 3.12 implies that a vertical vector field X_F is dissipative if and only if the corresponding force field f_H satisfies $\langle f_H, \dot{q} \rangle|_{(q,p)} \leq 0$ at all $(q, p) \in \mathbb{V}$.

DEFINITION 4.2. *A forced Hamiltonian system (2.3) is dissipative if the vector field can be decomposed as $X = X_H + X_F$, where X_H is a Hamiltonian vector field and X_F is a dissipative vector field.*

By (3.35), if the vertical vector field X_F is dissipative, then $\dot{E}(t) \leq 0$, which means that the system energy is nonincreasing in time.

In the last section, based on the empirical data of X_F , X_H , X , or f_H , we have discussed several approaches to extend the snapshot matrix such that the reduced model can quantitatively preserve the rate of energy dissipation. In the absence of the empirical data of vector fields, it is still desired for the reduced model to qualitatively preserve the *dissipativity*. This implies that if the original system is dissipative, then the reduced system should remain dissipative. Fortunately, when the dissipation is Rayleigh dissipation, the aforementioned structure-preserving projection automatically preserves the dissipativity, and this property is independent from the data that is used to construct the basis matrix A .

The dissipative force often arises from Rayleigh dissipation function, which can be written as

$$(4.1) \quad \mathcal{F}(q, \dot{q}) = \frac{1}{2} \dot{q}^T R(q) \dot{q}$$

in Lagrangian coordinates, where $R(q) \in \mathbb{R}^{n \times n}$ is a symmetric positive-semidefinite matrix. The force field is then given by $f_L(q, \dot{q}) = -\nabla_{\dot{q}} \mathcal{F}(q, \dot{q}) = -R(q)\dot{q}$. Using the Legendre transformation, we obtain $f_H(q, p) = -R(q)\dot{q}(q, p)$ in Hamiltonian coordinates. Since

$$\langle f_H, \dot{q} \rangle|_{(q,p)} = -\dot{q}^T R \dot{q}|_{(q,p)} \leq 0,$$

this verifies that the corresponding vertical vector field X_F is dissipative.

If the reduced system is constructed by the structure-preserving reduction, then $A_{qp} = 0$ and $q = A_{qq}r + A_{qp}s = A_{qq}r$. It follows that $\dot{q} = A_{qq}\dot{r}$. Using (3.7), the reduced force field is given by $A_{qq}^T f_H(r, s) = -A_{qq}^T R(A_{qq}r) A_{qq} \dot{r}$. Thus, the rate of energy variation of the reduced system at (r, s) is given by

$$\langle A_{qq}^T \tilde{f}_H, \dot{r} \rangle|_{(r,s)} = -\langle A_{qq}^T R(A_{qq}r) A_{qq} \dot{r}, \dot{r} \rangle|_{(r,s)} = -(A_{qq} \dot{r})^T R(A_{qq}r) (A_{qq} \dot{r})|_{(q,p)} \leq 0.$$

This verifies that the reduced system preserves the dissipativity.

Dissipativity preservation often is a strong indicator for stability preservation, as discussed in the next section.

4.2. Stability preservation. Let $\mathbb{V} = \mathbb{R}^{2n}$ be a configuration space with the standard topology induced by the Euclidean norm $\|\cdot\|$. Let M be a subset of \mathbb{V} . Then the subspace topology in M is the same as the metric topology obtained by restricting the Euclidean norm $\|\cdot\|$ to M . Since the Hamiltonian function $H : \mathbb{V} \rightarrow \mathbb{R}$ is continuous, the restriction of H to M gives a continuous function $H_M : M \rightarrow \mathbb{R}$. Throughout this section, we assume that the forced Hamiltonian system is dissipative, and the solution $x(t)$ of the system lie in M for every $t \geq 0$. We use $E_0 := H(x_0)$ to denote the system energy at $t = 0$.

Let $x(\mathbb{R}_+) = \{x(t) : t \geq 0\}$ denote the solution trajectory of a dissipative Hamiltonian system. We say the system is *uniformly bounded* if there exists a closed r -ball $B_r := \{x \in M : \|x\| \leq r\}$ centered at 0 such that $x(\mathbb{R}_+) \subset B_r$. Under certain conditions, the dissipative Hamiltonian system is uniformly bounded, as the following three lemmas indicate.

LEMMA 4.3. *Let $D := H_M^{-1}((-\infty, E_0]) = \{x \in M : H_M(x) \leq E_0\}$ denote a sublevel set of the Hamiltonian function $H_M : M \rightarrow \mathbb{R}$. Let D_0 be the connected component of D that contains x_0 . If D_0 is bounded, then the dissipative Hamiltonian system is uniformly bounded.*

Proof. Since $H \circ x : t \mapsto H_M(x(t))$ gives a continuous function of t , the set $x(\mathbb{R}_+) = \{x(t) : t \geq 0\}$ is path connected, hence it is connected. Since the forced Hamiltonian system is dissipative, $H_M(x(t)) \leq E_0$ for any $t \geq 0$. This implies that $x(\mathbb{R}_+) \subset D$. Since D_0 is a connected component of D and $D_0 \cap x(\mathbb{R}_+)$ contains x_0 , the connected set $x(\mathbb{R}_+)$ lies entirely within D_0 . Hence, if D_0 is bounded, so is $x(\mathbb{R}_+)$. \square

LEMMA 4.4. *If there exists a bounded neighborhood U of x_0 in M such that $E_0 < H_M(x)$ for every x on the boundary of U , then the dissipative Hamiltonian system is uniformly bounded.*

Proof. Let $\text{bd}_M(U)$ denote the boundary of U in M , and $\text{cl}_M(U)$ denote the closure of U in M . Since $H_M(x) > E_0$ for every $x \in \text{bd}_M(U)$, we have $D \subset M - \text{bd}_M(U)$. Since U and $M - \text{cl}_M(U)$ form a separation of $M - \text{bd}_M(U)$, as a connected set, D_0 must lie entirely within either U or $M - \text{cl}_M(U)$. Since $x_0 \in D_0 \cap U$, the only possible case is that $D_0 \subset U$. Because U is bounded, so is D_0 . By Lemma 4.3, the dissipative Hamiltonian system is uniformly bounded. \square

LEMMA 4.5. *If $\lim_{x \rightarrow \infty} H_M(x) = +\infty$ in M , then the dissipative Hamiltonian system is uniformly bounded.*

Proof. Suppose the system is not uniformly bounded. Then there exists an increasing sequence of time $\{t_1, t_2, \dots\}$ such that $\|x(t_i)\| > i$ for each $i \in \mathbb{N}_+$. By assumption, $H_M(x) \rightarrow +\infty$ as $x \rightarrow \infty$. Thus, for any $E_0 \in \mathbb{R}$, there exists an $n \in \mathbb{N}_+$ such that as long as $\|x\| > n$, $H_M(x) > E_0$. This implies that $H_M(x(t_i)) > E_0$ for every $i \geq n$. But if the system is dissipative, we must have $H_M(x(t_i)) \leq E_0$, which is a contradiction. \square

REMARK 4.6. If E_0 is a regular value of $H_M : M \rightarrow \mathbb{R}$, then the level set $H_M^{-1}(E_0)$ is an embedded codimension-1 submanifold in M by the regular value theorem, and the sublevel set D is an embedded codimension-0 submanifold with boundary in M [12] (pp. 120–121).

Let $M = \mathbb{V}$. Then, Lemmas 4.3, 4.4, and 4.5 imply that under certain conditions, the original dissipative Hamiltonian system is bounded. Moreover, if we respectively replace x_0 and E_0 by $x(t_1)$ and $H(x(t_1))$ for some $t_1 \in \mathbb{R}$, these lemmas still hold.

Next, we consider boundedness of the structure-preserving reduced model. Suppose that the reduced system remains dissipative, $x_0 \in \text{Range}(A)$, and the initial condition of the reduced system is given by $z_0 = A^+x_0$. Let $M = \text{Range}(A)$. Then, Lemmas 4.3, 4.4, and 4.5 imply that under the same conditions, the reduced dissipative Hamiltonian system preserves the boundedness. In particular, in Lemma 4.3, if the connected component D_0 of $H^{-1}((-\infty, E_0])$ in \mathbb{V} is bounded, then the connected component of $D_0 \cap M$ that contains x_0 is bounded in M . In Lemma 4.4, if there exists a bounded neighborhood U of x_0 in \mathbb{V} such that $E_0 < H(x)$ for every $x \in \text{bd}_{\mathbb{V}}(U)$, then $U_M := U \cap M$ is a neighborhood of x_0 in M and is bounded in M . Moreover, since $\text{bd}_M(U_M) \subset \text{bd}_{\mathbb{V}}(U) \cap M$, $E_0 < H(x)$ for every $x \in \text{bd}_M(U_M)$. In Lemma 4.5, if $\lim_{x \rightarrow \infty} H(x) = +\infty$ in \mathbb{V} , then $\lim_{x \rightarrow \infty} H_M(x) = +\infty$ in M .

Under the assumptions of Lemmas 4.3, 4.4, and 4.5, we have proved that the boundedness of the original and the reduced systems is consistent. In dynamical systems, boundedness is often accompanied with stability. An equilibrium point x_* of a dynamical system is *Lyapunov stable* if for every neighbourhood U of x_* , there exists a neighbourhood $V \subset U$ such that if $x_0 \in V$, then $x(t) \in U$ for every $t \geq 0$. When the system is linear and uniformly bounded, it is marginally stable in the sense of Lyapunov. If the original forced Hamiltonian system is linear, then the reduced system constructed by the structure-preserving projection is also linear. Thus, if any assumption in the previous lemmas holds, both the original and reduced systems are Lyapunov stable.

THEOREM 4.7. *Let M be a closed subset of \mathbb{V} . If x_* is a strict local minimum of H_M in M , then x_* is a stable equilibrium for the dissipative Hamiltonian system.*

Proof. Since x_* is a strict local minimum of H_M , then there exists a neighbourhood W of x_* such that $H_M(x) > H_M(x_*)$ for every $x \in \text{cl}_M(W) - \{x_*\}$. Assume W is bounded in M , otherwise, replace W by $W \cap B_r(x_*)$ for an open r -ball $B_r(x_*)$ centered at x_* . Let U be an arbitrary neighborhood of x_* in M . Since both W and U are open, so is $W \cap U$. Let $U_0 = W \cap U$. Since $\text{bd}_M(U_0) \subset \text{cl}_M(W)$, $\text{bd}_M(U_0)$ is bounded in M , and hence also bounded in \mathbb{V} . Since M is closed in \mathbb{V} , $\text{bd}_M(U_0)$ is also closed in \mathbb{V} . As a bounded and closed subset of \mathbb{V} , $\text{bd}_M(U_0)$ is compact. By the extreme value theorem, there exists $x_1 \in \text{bd}_M(U_0)$ such that $H_M(x_1) \leq H_M(x)$ for every $x \in \text{bd}_M(U_0)$. Since H_M is continuous, the preimage $D := H_M^{-1}((-\infty, H_M(x_1)))$ of $(-\infty, H(x_1))$

TABLE 4.1

The POD-Galerkin method vs. the structure-preserving model reduction method.

	POD-Galerkin	Structure-preserving model reduction	
Original system	General ODE system: $\dot{x} = f(x)$ with $x \in \mathbb{R}^n$	Forced Hamiltonian system: $\dot{x} = X_H(x) + X_F(x)$ with $x \in \mathbb{R}^{2n}$	
Physical laws of the original system	Newton's Law	Integral d'Alembert's principle	Local d'Alembert's principle
Reduced state	Orthogonal projection: $z = \Phi^T x \in \mathbb{R}^k$	Symplectic projection: $z = A^+ x \in \mathbb{R}^{2k}$	
Reduced system	Reduced ODE system: $\dot{z} = \Phi^T f(\Phi z)$	Reduced forced Hamiltonian system: $\dot{z} = X_{\tilde{H}}(z) + X_{\tilde{F}}(z)$	
Reduction approach	Galerkin projection	Variational principle	Structure-preserving projection
Physical laws of the reduced system	N/A	Integral d'Alembert's principle	Local d'Alembert's principle
Basis matrix	Orthonormal: $\Phi^T \Phi = I_k$	Symplectic: $A^T J_{2n} A = J_{2k}$ and $A_{qp} = 0$	
Basis matrix construction method	POD: SVD	PSD: Cotangent lift	
Dissipativity	N/A	Dissipativity preservation	
Stability	N/A	Stability preservation	

is open in M . Since $x_1 \in \text{cl}_M(W)$ and $x_1 \neq x_*$, we have $H_M(x_*) < H_M(x_1)$, which implies that $x_* \in D$. Thus, $V := U_0 \cap D$ is a neighbourhood of x_* in M . If $x_0 \in V$, then $H_M(x_0) < H_M(x_1) \leq H_M(x)$ for every $x \in \text{bd}_M(U_0)$. This implies that $x(\mathbb{R}_+) \subset U_0 \subset U$, by Lemma 4.4. Therefore, x_* is a stable equilibrium for the dissipative Hamiltonian system. \square

Let $M = \mathbb{V}$. Suppose U is a neighborhood of x_* in \mathbb{V} , and x_* is the minimum of H in U . Then, Theorem 4.7 implies that the full model is stable at x_* . Now, let $M = \text{Range}(A)$. It immediately follows that x_* is also the minimum of H_M in U_M , where $U_M = U \cap M$. It follows that x_* is also the stable equilibrium of the reduced Hamiltonian system on $\text{Range}(A)$. Therefore, the stability of the full and reduced dissipative Hamiltonian systems is consistent. For both the full model and the structure-preserving reduced model, H_M can be considered a Lyapunov function for the system. Nevertheless, a POD reduced system is not guaranteed to be dissipative and stable, and therefore, there is no corresponding Lyapunov function.

While both the structure-preserving method and the POD-Galerkin method construct reduced equations in some low dimensional subspaces, only the structure-preserving method can preserve the forced-Hamiltonian structure. The PSD algorithm can be used to construct a symplectic matrix A , which is an analogous to POD that constructs an orthonormal basis matrix Φ . Evolving a PSD reduced system by a symplectic integrator can capture the energy variation and preserve the stability. By contrast, even if a POD subspace can fit the empirical data with good accuracy, a POD reduced system can be unstable. To this end, one can distinguish between a numerically reduced system and a physically reduced system. Table 4.1 compares the POD-Galerkin method with the proposed structure-preserving model reduction method; it serves as a short summary of sections 3–4.

5. Numerical validation. In this section, the performance of the proposed structure-preserving model reduction method is illustrated in numerical simulation of a linear dissipative wave equation. Our goal is to demonstrate that PSD can deliver a low-dimensional reduced system while preserving the stability of the original system.

5.1. Hamiltonian formulation of dissipative wave equations. Let $u = u(t, x)$. Consider a one-dimensional linear wave equation with constant damping coefficient β , undamped angular frequency ω_0 , and moving speed c ,

$$(5.1) \quad u_{tt} + \beta u_t - c^2 u_{xx} + \omega_0^2 u = 0,$$

on space $x \in [0, l]$. With the generalized coordinates $q = u$ and the generalized momenta $p = u_t$, the Hamiltonian PDE associated with (5.1) is given by

$$(5.2) \quad \dot{q} = \frac{\delta H}{\delta p}, \quad \dot{p} = -\frac{\delta H}{\delta q} - \beta p,$$

where the Hamiltonian is defined as

$$(5.3) \quad H(q, p) = \int_0^l dx \left[\frac{1}{2} p^2 + \frac{1}{2} \omega_0^2 q^2 + \frac{1}{2} c^2 q_x^2 \right].$$

A fully resolved model of (5.2) can be constructed by a structure-preserving finite difference discretization [2]. In particular, with n equally spaced grid points, the spatially discretized Hamiltonian with periodic boundary conditions is given by

$$(5.4) \quad H_d(y) = \frac{\Delta x}{2} \sum_{i=1}^n p_i^2 + \frac{\omega_0^2 \Delta x}{2} \sum_{i=1}^n q_i^2 + \frac{c^2}{2\Delta x} \sum_{i=1}^n (q_i - q_{i-1})^2,$$

where $x_i = i\Delta x$, $q_i = u(t, x_i)$, $q_0 = q_n$, $p_i = u_t(t, x_i)$, and $y = [q_1; \dots; q_n; p_1; \dots; p_n]$. With $n\Delta x = l$, (5.4) converges to (5.3) in the limit $\Delta x \rightarrow 0$. Now, we have a Hamiltonian ODE system,

$$(5.5) \quad \frac{dy}{dt} = J_d \nabla_y H_d + X_F,$$

where $J_d = J_{2n}/\Delta x$, and $X_F = [0; \dots; 0; -\beta p_1; \dots; -\beta p_n]$. Let $D_{xx} \in \mathbb{R}^{n \times n}$ denote the three-point central difference approximation for the spatial derivative ∂_{xx} . We define a Hamiltonian matrix K and a dissipative matrix L by

$$(5.6) \quad K = \begin{bmatrix} 0_n & I_n \\ c^2 D_{xx} - \omega_0^2 I_n & 0_n \end{bmatrix}, \quad L = \begin{bmatrix} 0_n & 0_n \\ 0_n & -\beta I_n \end{bmatrix}.$$

Then, (5.5) can be written in the form

$$(5.7) \quad \dot{y} = Ky + Ly.$$

Time discretization of (5.7) can be achieved by using an implicit symplectic integrator scheme based on mid-point rule [6, 15].

5.2. Numerical results. For our numerical experiments, we study the one-dimensional dissipative wave equation with periodic boundary conditions defined in (5.1). Let $s(x) = 10 \times |x - \frac{1}{2}|$; and let $h(s)$ be a cubic spline function:

$$h(s) = \begin{cases} 1 - \frac{3}{2}s^2 + \frac{3}{4}s^3 & \text{if } 0 \leq s \leq 1 \\ \frac{1}{4}(2-s)^3 & \text{if } 1 < s \leq 2 \\ 0 & \text{if } s > 2 \end{cases}.$$

The initial condition is provided by

$$(5.8) \quad q(0) = [h(s(x_1)); \dots; h(s(x_n))], \quad p(0) = 0_{n \times 1},$$

which gives rise to a dissipative system with wave propagating in both directions of x and then bouncing back. The full model is computed using the following parameter set:

Size of the space domain	$l = 1$
Number of grid points	$n = 500$
Space discretization step	$\Delta x = l/n = 0.002$
Final time	$T = 50$
Time discretization step	$\delta t = 0.01$
Damping coefficient	$\beta = 0.1$
Undamped angular frequency	$\omega_0 = 0.05$
Wave speed	$c = 0.1$

The reduced PSD model is constructed through the cotangent lift method based on the extended snapshot matrix (3.32) that contains snapshots of $q(t)$ and $p(t)$. Since $f_H = -\beta p$, this extended snapshot matrix can also discover the dominant modes of f_H , and therefore approximately preserves the system energy. Since (5.1) is linear, we can also obtain the analytical solution by the eigenfunction expansion method. The analytical solution is used as the reference benchmark solver to measure the error of the full model as well as POD and PSD reduced models.

Figure 5.1(a) plots several snapshots of the solution profile from $t = 0$ to $t = 10$. The empirical data ensemble takes 101 snapshots from the full model with uniform interval ($\Delta t = 0.5$). We first compare PSD with the full model. The lines show the results from the full model and the symbols show the results from the PSD reduced model with 20 modes. For all snapshots, the PSD reduced system obtains good results that match the full model very well. Figure 5.1(b) shows the singular values corresponding to the first 80 POD and PSD modes. A fast decay of singular values indicates a fast convergence of low-dimensional data to fit the original data with respect to the L^2 norm. Since POD is designed to minimize the projection error of the data snapshots in least-squares sense, for a fixed dimension, no other linear projection method can provide better data approximation with the L^2 norm. With the symplectic constraint, we do observe that the cotangent lift requires more modes to fit the empirical data than POD in order to obtain the same accuracy.

However, preserving the data does not necessarily imply preserving the dynamics. With more modes, there is no guarantee that the POD reduced system will yield more accurate solutions. As Figure 5.2 indicates, the L^2 error norm of the POD reduced system increases exponentially when it has 20, 30, or 40 modes. In addition, the POD reduced system with 40 modes blows up faster than the POD system with 20 modes. This result verifies that POD can yield unstable reduced systems, even though the original system is dissipative and stable. By contrast, PSD reduced systems have small numerical errors in the L^2 norm for all the tested cases. Figure 5.3 shows that PSD reduced models accurately capture the evolution of the system energy E of the dissipative wave equation, while the energy of POD reduced systems quickly grows to infinity. In this example, increasing the number of POD modes actually causes the system energy to increase at a faster rate. Here, E equals the discretized Hamiltonian $H_d(y)$.

Figure 5.4(a) plots the L^2 norm of the total error of different systems over the whole time domain $[0, 50]$. We compare the full model (with $k = 1000$), coarse model,

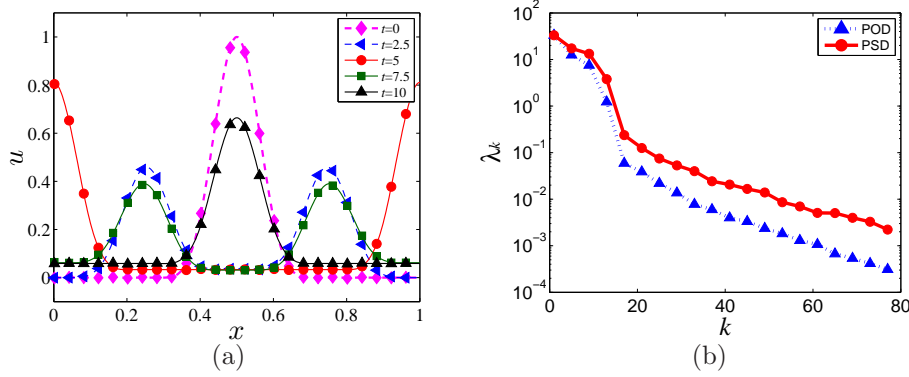


FIGURE 5.1. (Color online.) (a) The solution $u(t, x)$ at $t = 0, 2.5, 5, 7.5, 10$ of the dissipative wave equation. The lines represent the results from the full model based on 500 grid points and the symbols represent the results from the PSD reduced model with 20 modes. (b) The singular values λ_k corresponding to the first $k = 1, 2, \dots, 80$ POD and PSD modes.

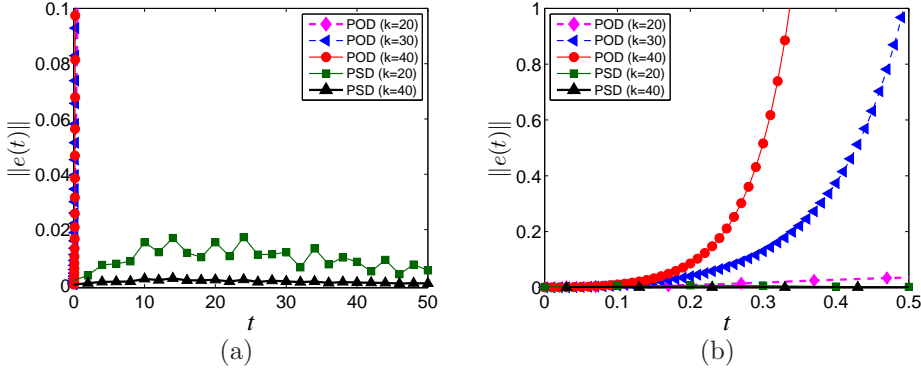


FIGURE 5.2. (Color online.) Comparison between POD and PSD (cotangent lift) reduced systems of the dissipative wave equation. (a) The evolution of the L^2 error norm, $\|e(t)\| := \|\hat{u}(t) - u(t)\|$, between the benchmark solution $u(t)$ and approximating solutions $\hat{u}(t)$ for the time domain $t \in [0, 50]$. (b) The L^2 error norm $\|e(t)\|$ for the zoomed in time interval $t \in [0, 0.5]$.

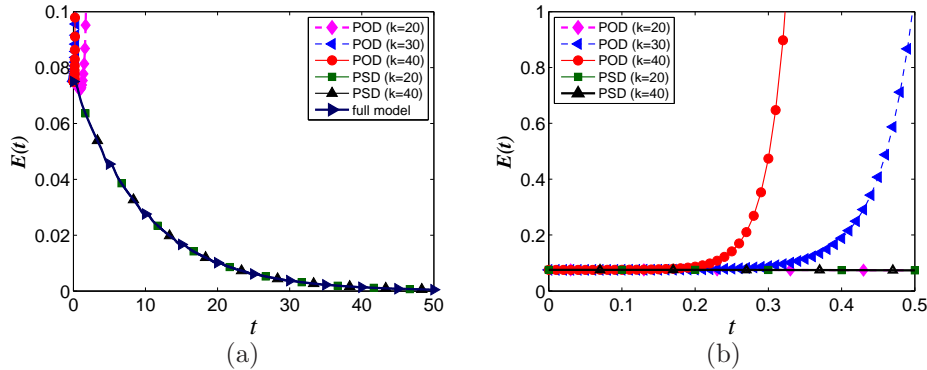


FIGURE 5.3. (Color online.) Comparison between full model, POD reduced system, and PSD (cotangent lift) reduced system of the dissipative wave equation. (a) The evolution of the system energy $E(t)$ for the time domain $t \in [0, 50]$. (b) The system energy $E(t)$ for the zoomed in time interval $t \in [0, 0.5]$.

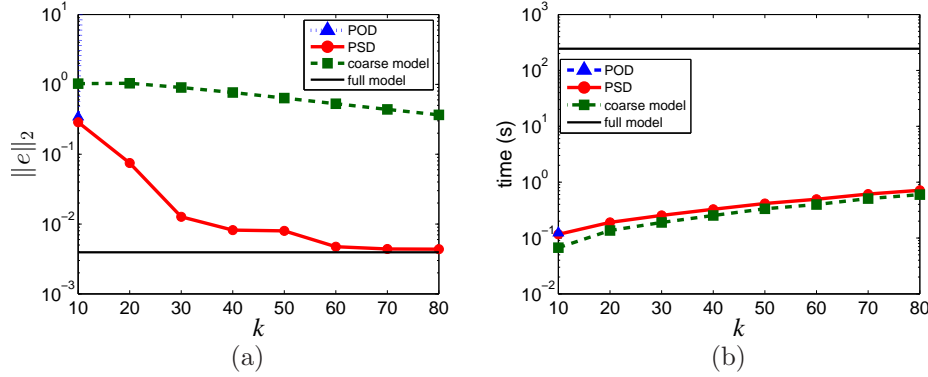


FIGURE 5.4. (Color online.) Comparison between the full model, coarse model, POD reduced system, and PSD (cotangent lift) reduced system for the dissipative wave equation. (a) The L^2 norm of the total error $\|e\|_2 := \sqrt{\int_0^T \|e(t)\|^2 dt}$ of different systems. For the POD reduced system, we only compute $\|e\|_2$ with 10 modes; when the subspace dimension k is greater than 10, the POD reduced system blows up in the whole time domain $[0, 50]$ and $\|e\|_2$ becomes infinite. (b) The running time of different systems corresponding to different subspace dimensions k . All data comes from the average value of 10 independent runs.

as well as POD and PSD reduced model. The subspace dimension k of the coarse model and reduced models ranges from 10 to 80. The L^2 norm of the total error of the POD reduced system is bounded only when k equals 10 for all the tested cases. While the PSD reduced system shows some numerical error, this error quickly converges to the error of the full model. The coarse model also preserves the forced Hamiltonian structure and remains stable, but the numerical error of the coarse model reduces at a low rate with increased modes. Figure 5.4(b) shows the running time of different methods. We find coarse model and POD/PSD reduced model have similar running speed. With 80 modes, both the coarse model and the reduced model can significantly improve the computational efficiency and reduce the running time by more than two orders of magnitude.

5.3. Stability analysis. Using the numerical results, we further analyze the stability for the linear system in (5.1). Using (5.4), we know $\lim_{x \rightarrow \infty} H_d(x) = +\infty$; Lemma 4.5 implies that the full model is uniformly bounded. Since the origin is the strict minimum of H_d , Theorem 4.7 implies that the origin is a stable equilibrium for the dissipative wave equation. Since the external force of (5.1) is a Rayleigh dissipative force, where $\mathcal{F}(q, \dot{q}) = \frac{1}{2} \dot{q}^T \dot{q}$, the reduced PSD system is also dissipative. By the same argument, the reduced PSD system is uniformly bounded, and also has the origin as a stable equilibrium.

To explain why the POD reduced system is unstable, we study the eigenvalues of the linear wave equation. According to [13], the eigenvalues β_i ($i = 1, \dots, n$) of the discretized spatial derivative D_{xx} with periodic boundary conditions are given by

$$\beta_i = -\frac{2}{\Delta x^2} \left[1 - \cos\left(\frac{2\pi i}{n}\right) \right].$$

It follows that the eigenvalues of the full model $K + L$ in (5.7) are given by $2n$ complex numbers $\{\lambda_i\}_{i=1}^{2n}$, where λ_i, λ_{i+n} are solutions of $\lambda^2 + \beta\lambda - c^2\beta_i + \omega_0^2 = 0$ for $i = 1, \dots, n$. It can be verified that all the eigenvalues of the full model have negative real parts, which means the full model is stable.

Since POD does not preserve the system energy, there are no mechanisms to confine the solution in a bounded region. As a result, the reduced system may blow up with time evolution. To corroborate this claim, let Φ denote a POD basis matrix, λ_* denote the eigenvalue of $\Phi^T(K + L)\Phi$ with the maximal real part, and ξ_* denote the corresponding eigenvector with unit length. Then, $a_* = \xi_*^T y_0$ gives the projection coefficient of y_0 onto ξ_* . Since the solution of a linear system has an exponential term $a_* \exp(\lambda_* t) \xi_*$, the POD reduced system is unstable when $a_* \neq 0$ and $\text{Re}(\lambda_*) > 0$.

Table 5.1 lists $\text{Re}(\lambda_*)$ with a wide range of diffusion coefficients β and subspace dimensions k . Numerical results show that $a_* \neq 0$ for all the tested cases. When $\beta = 10^2$, the diffusion term becomes dominant in (5.7). The POD reduced system is stable when $k = 10$ and $k = 20$ for the tested cases. When $10^{-2} \leq \beta \leq 10^1$, The POD reduced system is stable only when $k = 10$. When $\beta = 10^{-3}$, the diffusion term becomes negligible in (5.7) and the POD reduced system is unstable for all the tested cases. Table 5.1 also shows that when $\beta = 10^{-1}$, $\text{Re}(\lambda_*)$ with 40 modes is much larger than $\text{Re}(\lambda_*)$ with 20 modes, which explains why the POD reduced system with 40 modes blows up faster than the system with 20 modes in Figure 5.2.

TABLE 5.1

The real part $\text{Re}(\lambda_)$ of the eigenvalue corresponding to the most unstable POD mode for different diffusion coefficients β and subspace dimensions k .*

β	k							
	10	20	30	40	50	60	70	80
10^{-3}	5.17×10^{-3}	0.304	15.1	19.9	16.7	17.4	19.6	111
10^{-2}	-3.71×10^{-3}	0.252	15.6	20.2	17.0	17.9	19.7	113
10^{-1}	-2.09×10^{-2}	1.26	12.3	18.0	21.8	20.5	22.3	129
10^0	-2.51×10^{-3}	1.43	31.4	37.2	26.3	60.3	44.6	139
10^1	-2.50×10^{-4}	1.08	18.7	26.7	38.0	43.1	47.8	50.0
10^2	-1.07×10^{-3}	-2.50×10^{-5}	3.66	32.8	33.6	43.8	54.7	64.2

6. Conclusion. This paper proposed a PSD model reduction method to simplify large-scale forced Hamiltonian systems, which can achieve significant computational savings. Since the PSD reduced system preserves the forced Hamiltonian structure, it automatically satisfies the d'Alembert's principle. Since d'Alembert's principle is the first principle in classical mechanics, the PSD reduced system is a physical model, rather than merely a numerical model. In contrast, although POD can always reduce the dimensionality of a dynamical system, a POD reduced system may be or may not be physical, since there is no guarantee that the system can satisfy any fundamental physical laws.

Two structure-preserving approaches are developed in order to reconstruct reduced systems in a low-dimensional subspace, one based on the variational principle and the other on the structure-preserving projection. Both approaches can yield the same structure-preserving reduced system. By incorporating the vector field into the data ensemble, the PSD method also preserve the time derivative of system energy. In a special case when the external force represents the Rayleigh dissipation, PSD automatically preserves the dissipativity. As a consequence, PSD also preserves the boundedness and Lyapunov stability under some conditions.

The stability, accuracy, and efficiency of the proposed method are illustrated through numerical simulations of the one dimensional dissipative wave equation. However, PSD can have much more general applications. Once we choose canonical coor-

dinates, all the systems that satisfy d'Alembert's principle can be written as the forced Hamiltonian equation. As a result, PSD can be applied to any large-scale mechanical system in principle. Finally, we should mention that the computational complexity and implementation complexity of PSD are almost identical to the complexity of POD. Since the POD reduced system can be unstable and produces unpredictable results, we believe that PSD is more suited for model reduction of large-scale mechanical systems, especially when long-time integration is required.

REFERENCES

- [1] A. C. ANTOUNAS, D. C. SORESENSEN, AND S. GUGERCIN, *A survey of model reduction methods for large-scale systems*, Contemp. Math., 280 (2001), pp. 193–219.
- [2] T. J. BRIDGES AND S. REICH, *Numerical methods for Hamiltonian PDEs*, J. Phys. A, 39 (2006), pp. 5287–5320.
- [3] K. CARLBERG AND C. FARHAT, *A low-cost, goal-oriented 'compact proper orthogonal decomposition' basis for model reduction of static systems*, Internat. J. Numer. Methods Engrg., 86 (2011), pp. 381–402.
- [4] K. CARLBERG, R. TUMINARO, AND P. BOGGS, *Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics*, SIAM J. Sci. Comput., 37 (2015), pp. B153–B184.
- [5] S. GUGERCIN, R. V. POLYUGA, C. BEATTIE, AND A. J. VAN DER SCHAFT, *Structure-preserving tangential interpolation for model reduction of port-Hamiltonian systems*, Automatica, 48 (2012), pp. 1963–1974.
- [6] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer Ser. Comput. Math. 31, Springer-Verlag, Berlin, 2006.
- [7] C. HARTMANN, V.-M. VULCANOV, AND C. SCHÜTTE, *Balanced truncation of linear second-order systems: A Hamiltonian approach*, SIAM J. Multiscale Model. and Simul., 8 (2010), pp. 1348–1367.
- [8] T. ILIESCU AND Z. WANG, *Are the snapshot difference quotients needed in the proper orthogonal decomposition?*, SIAM J. Sci. Comput., 36 (2014), pp. A1221–A1250.
- [9] K. KUNISCH AND S. VOLKWEIN, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Numer. Math., 90 (2001), pp. 117–148.
- [10] ———, *Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics*, SIAM J. Numer. Anal., 40 (2002), pp. 492–515.
- [11] S. LALL, P. KRYSL, AND J. E. MARSDEN, *Structure-preserving model reduction for mechanical systems*, Phys. D, 184 (2003), pp. 304–318.
- [12] J. M. LEE, *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics (Book 218), Springer, 2nd ed., 2012.
- [13] R. J. LEVEQUE, *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*, SIAM, Philadelphia, 2007.
- [14] J. E. MARSDEN AND T. RATIU, *Introduction to Mechanics and Symmetry*, 2nd ed., Texts in Appl. Math. 17, Springer-Verlag, New York, 2003.
- [15] R. I. MCLACHLAN AND G. R. W. QUISPTEL, *Geometric integrators for ODEs*, J. Phys. A, 39 (2006), pp. 5251–5285.
- [16] B. C. MOORE, *Principal component analysis in linear systems: Controllability, observability, and model reduction*, IEEE Trans. Automat. Control, 26 (1981), pp. 17–32.
- [17] L. PENG AND K. MOHSENI, *An online manifold learning approach for model reduction of dynamical systems*, SIAM J. Numer. Anal., 52 (2014), pp. 1928–1952.
- [18] ———, *Symplectic model reduction of Hamiltonian systems*, SIAM J. Sci. Comput., 38 (2016), pp. A1–A27.
- [19] R. V. POLYUGA AND A. J. VAN DER SCHAFT, *Structure preserving model reduction of port-Hamiltonian systems by moment matching at infinity*, Automatica, 46 (2010), pp. 665–672.
- [20] S. PRAJNA, *POD model reduction with stability guarantee*, in Proceedings of the 42nd IEEE Conference on Decision and Control, vol. 5, Maui, HI, 2003, pp. 5254–5258.
- [21] M. RATHINAM AND L. R. PETZOLD, *A new look at proper orthogonal decomposition*, SIAM J. Numer. Anal., 41 (2003), pp. 1893–1925.
- [22] L. N. TREFETHEN AND D. BAU, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.